

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
17 May 2001 (17.05.2001)

PCT

(10) International Publication Number  
**WO 01/34809 A2**

(51) International Patent Classification<sup>7</sup>: **C12N 15/31**,  
C07K 14/31, 16/12, C12Q 1/68, G01N 33/53, A61K  
39/085, 48/00

(74) Agents: **LEVY, David, J.**; Glaxo Wellcome Inc., Five  
Moore Drive, P.O. Box 13398, Research Triangle Park, NC  
27709-3398 et al. (US).

(21) International Application Number: **PCT/US00/30782**

(22) International Filing Date:  
9 November 2000 (09.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/164,258 9 November 1999 (09.11.1999) US

(71) Applicant (for all designated States except US): **GLAXO  
GROUP LIMITED** [GB/GB]; Glaxo Wellcome House,  
Berkeley Avenue, Greenford, Middlesex UB6 0NN (GB).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **KIMMERLY,  
William, John** [US/US]; 242 Via Villena, Encinitas, CA  
92024 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— Without international search report and to be republished  
upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

WO 01/34809 A2

(54) Title: **STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS**

(57) Abstract: S epidermidis polypeptides and DNA (RNA) encoding such polypeptides and a procedure for producing such polypeptides by recombinant techniques is disclosed. Also disclosed are methods for utilizing such polypeptides and DNA (RNA) for the treatment of infection, particularly infections arising from S epidermidis. Antagonists against the function of such polypeptides and their use as therapeutics to treat infection are also disclosed. Also disclosed are diagnostic assays for detecting diseases related to the presence of S epidermidis nucleic acid sequences and the polypeptides in a host. Also disclosed are diagnostic assays for detecting polynucleotides and polypeptides related to S epidermidis.

## STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

### Field of the Invention

The present invention provides nucleic acids, and peptides, polypeptides and proteins encoded by the nucleic acids, isolated from *Staphylococcus*  
5 *epidermidis*.

### Background of the Invention

*Staphylococcus epidermidis* is a gram-positive bacteria present in the normal flora of humans, and is typically present on the skin. It is catalase positive  
10 and grows aerobically. It is implicated in various human conditions and diseases, including subacute bacterial endocarditis (Baddour LM et al., Production of experimental endocarditis by coagulase-negative staphylococci: variability in species virulence, J. Infect. Dis. 150: 721-727, 1984; Karchmer AW, Archer GL, Dismukes WE, *Staphylococcus epidermidis* causing prosthetic valve endocarditis:  
15 microbiologic and clinical observations as guides to therapy, Ann Intern Med. 1983;98:447-455.) and septicemia (Christensen GD et al., Nosocomial septicemia due to multiply antibiotic-resistant *Staphylococcus epidermidis*, Ann. Intern. Med. 96: 1-10, 1982). *S. epidermidis* is estimated to be responsible for about 12% of all hospital patient infections. Because of the organism's peculiar ability to colonize  
20 polymer and metallic surfaces, there is a correlation of infection with the insertion of intravenous lines or catheters or implantation of prosthetic devices. Treatment can be difficult since different isolates of *S. epidermidis* show a broad spectrum of antibiotic resistance. The organism also produces a polysaccharide biofilm which helps to protect the bacteria from the human immune system (Tojo M et al.,  
25 Isolation and characterization of a capsular polysaccharide adhesin from *Staphylococcus epidermidis*, J. Infect. Dis. 157: 713-722, 1988).

The present invention advantageously provides isolated nucleic acids and their encoded peptides, polypeptides and proteins from the genome of *S. epidermidis*, as well as the genomic map of *S. epidermidis*. Thus, the present  
30 invention fulfils a a widely-felt need for *S. epidermidis* diagnostics, antigens, and

products useful in procedures for preparing antibodies and for identifying compounds effective against *S. epidermidis* infection. Selected nucleic acids and/or polypeptides of the present invention can be advantageously utilized as targets in screenings assays for antibiotics, as diagnostics of infections, and as means to identify *S. epidermidis* in any given sample and distinguish it from other bacteria.

### SUMMARY OF THE INVENTION

The present invention provides an isolated polynucleotide comprising a member selected from the group consisting of:

- (a) a polynucleotide encoding a polypeptide having at least a 70% identity to a polypeptide set forth in the Sequence Listing;
- (b) a polynucleotide which is complementary to the polynucleotide of (a); and
- (c) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a) or (b). The present invention further provides polypeptides encoded by these polynucleotides and methods of using the polynucleotides and polypeptides..

### DETAILED DESCRIPTION OF THE INVENTION

#### GLOSSARY

The following illustrative explanations are provided to facilitate understanding of certain terms used frequently herein, particularly in the Examples. The explanations are provided as a convenience and are not limitative of the invention.

**BINDING MOLECULE** refers to a molecule or ion which binds or interacts specifically with polypeptides or polynucleotides of the present invention, including, for example enzyme substrates, cell membrane components and classical receptors. Binding between polypeptides (or polynucleotides) of the invention and such molecules may be exclusive to polypeptides of the invention, which is preferred, or it may be highly specific for polypeptides of the invention, which is also preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes a polypeptide of the invention. Binding molecules also include antibodies and antibody-derived reagents that bind specifically to polypeptides of the invention.

GENETIC ELEMENT generally means a polynucleotide comprising a region that encodes a polypeptide or a polynucleotide region that regulates replication, transcription or translation or other processes important to expression of the polypeptide in a host cell, or a polynucleotide comprising both a region that encodes a polypeptide and a region operably linked thereto that regulates expression. Genetic elements may be comprised within a vector that replicates as an episomal element; that is, as a molecule physically independent of the host cell genome. They may be comprised within plasmids. Genetic elements also may be comprised within a host cell genome; not in their natural state but, rather, following manipulation such as isolation, cloning and introduction into a host cell in the form of purified DNA or in a vector, among others.

HOST CELL is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

IDENTITY, as known in the art, is the relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, identity also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. Identity can be readily calculated (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or two polypeptide sequences, the term is well known to skilled artisans (*Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988)). Methods commonly employed to determine identity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48:1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity are codified in computer programs. Preferred computer

program methods to determine identity between two sequences include, but are not limited to, GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol.* 215: 403 (1990)).

5           ISOLATED means separated "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living organism in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials  
10 of its natural state is "isolated", as the term is employed herein. As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in  
15 whole organisms. Introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as a media formulations, solutions for introduction of polynucleotides or polypeptides, for example, into cells, compositions  
20 or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

POLYNUCLEOTIDE(S) generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or  
25 DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more  
30 typically, double-stranded, or triple-stranded, or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a  
35 region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term polynucleotide includes DNAs or

5 RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA  
10 characteristic of viruses and cells, including simple and complex cells, *inter alia*. The term polynucleotide also embraces short polynucleotides often referred to as oligonucleotide(s). "Polynucleotide" and "nucleic acid" are often used interchangeably herein..

POLYPEPTIDES, as used herein, includes all polypeptides as described  
15 below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, unless otherwise indicated, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to  
20 longer chains, which generally are referred to in the art as proteins, of which there are many types. It will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and  
25 other post-translational modifications, but also by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic texts and in more detailed monographs, as well as in a  
30 voluminous research literature, and they are well known to those of skill in the art. Among the known modifications which may be present in polypeptides of the present are, to name an illustrative few, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or  
35 lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation

of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance

10 *PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., *Posttranslational Protein Modifications: Perspectives and Prospects*, pgs. 1-12 in *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS* B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., *Meth. Enzymol.* 182:626-646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, Ann. N.Y. Acad. Sci. 663: 48-62 (1992). It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be generally as a result of

20 posttranslational events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in

25 polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli* or other cells, prior to proteolytic processing, almost invariably will be N-formylmethionine. During post-translational modification of the peptide, a methionine residue at the NH<sub>2</sub>-terminus may be deleted. Accordingly, this invention contemplates the use of both the methionine-containing and the methionineless amino terminal variants of the protein of the invention. The modifications that occur in a polypeptide often will be a function of how it is

35 made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host

cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as, for example, *E. coli*. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cell often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation, *inter alia*. Similar considerations apply to other modifications. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized recombinantly by expressing a polynucleotide in a host cell.

VARIANT(S) of polynucleotides or polypeptides, as the term is used herein, are polynucleotides or polypeptides that differ from a reference polynucleotide or polypeptide, respectively. Variants in this sense are described below and elsewhere in the present disclosure in greater detail. (1) A polynucleotide that differs in nucleotide sequence from another, reference polynucleotide. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical. As noted below, changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide with the same amino acid sequence as the reference. Also as noted below, changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. (2) A polypeptide that differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.



Techniques are available to evaluate temporal gene expression in bacteria, particularly as it applies to viability under laboratory and host infection conditions. A number of methods can be used to identify genes which are essential to survival *per se*, or essential to the establishment/maintenance of an infection. Identification of expression of a sequence by one of these methods yields additional information about its function and permits the selection of such sequence for further development as a screening target. Briefly, these approaches include:

1) Signature Tagged Mutagenesis (STM)

This technique is described by Hensel *et al.*, *Science* 269: 400-403(1995), the contents of which is incorporated by reference for background purposes. Signature tagged mutagenesis identifies genes necessary for the establishment/maintenance of infection in a given infection model.

The basis of the technique is the random mutagenesis of target organism by various means (e.g., transposons) such that unique DNA sequence tags are inserted in close proximity to the site of mutation. The tags from a mixed population of bacterial mutants and bacteria recovered from an infected host are detected by amplification, radiolabeling and hybridization analysis. Mutants attenuated in virulence are revealed by absence of the tag from the pool of bacteria recovered from infected hosts.

In *Streptococcus pneumoniae*, because the transposon system is less well developed, a more efficient way of creating the tagged mutants is to use the insertion-duplication mutagenesis technique as described by Morrison *et al.*, *J. Bacteriol.* 159:870 (1984) the contents of which is incorporated by reference for background purposes.

2) In Vivo Expression Technology (IVET)

This technique is described by Camilli *et al.*, *Proc. Nat'l. Acad. Sci. USA.* 91:2634-2638 (1994), the contents of which is incorporated by reference for background purposes. IVET identifies genes up-regulated during infection when compared to laboratory cultivation, implying an important role in infection. Sequences identified by this technique are implied to have a significant role in infection establishment/maintenance.

In this technique random chromosomal fragments of target organism are cloned upstream of a promoter-less reporter gene in a plasmid vector. The pool is introduced into a host and at various times after infection bacteria may be recovered and assessed for the presence of reporter gene expression. The

chromosomal fragment carried upstream of an expressed reporter gene should carry a promoter or portion of a gene normally upregulated during infection. Sequencing upstream of the reporter gene allows identification of the up regulated gene.

5 3) Differential display

This technique is described by Chuang *et al.*, *J. Bacteriol.* 175:2026-2036 (1993), the contents of which is incorporated by reference for background purposes. This method identifies those genes which are expressed in an organism by identifying mRNA present using randomly-primed RT-PCR. By  
10 comparing pre-infection and post infection profiles, genes up and down regulated during infection can be identified and the RT-PCR product sequenced and matched to library sequences.

4) Generation of conditional lethal mutants by transposon mutagenesis.

This technique, described by de Lorenzo, V. *et al.*, *Gene* 123:17-24  
15 (1993); Neuwald, A. F. *et al.*, *Gene* 125: 69-73(1993); and Takiff, H. E. *et al.*, *J. Bacteriol.* 174:1544-1553(1992), the contents of which is incorporated by reference for background purposes, identifies genes whose expression are essential for cell viability.

In this technique transposons carrying controllable promoters, which  
20 provide transcription outward from the transposon in one or both directions, are generated. Random insertion of these transposons into target organisms and subsequent isolation of insertion mutants in the presence of inducer of promoter activity ensures that insertions which separate promoter from coding region of a gene whose expression is essential for cell viability will be recovered.  
25 Subsequent replica plating in the absence of inducer identifies such insertions, since they fail to survive. Sequencing of the flanking regions of the transposon allows identification of site of insertion and identification of the gene disrupted. Close monitoring of the changes in cellular processes/morphology during growth in the absence of inducer yields information on likely function of the gene. Such  
30 monitoring could include flow cytometry (cell division, lysis, redox potential, DNA replication), incorporation of radiochemically labeled precursors into DNA, RNA, protein, lipid, peptidoglycan, monitoring reporter enzyme gene fusions which respond to known cellular stresses.

5) Generation of conditional lethal mutants by chemical mutagenesis.

35 This technique is described by Beckwith, J., *Methods in Enzymology* 204:

3-18(1991), the contents of which are incorporated herein by reference for background purposes. In this technique random chemical mutagenesis of target organism, growth at temperature other than physiological temperature (permissive temperature) and subsequent replica plating and growth at different temperature (e.g. 42°C to identify ts, 25°C to identify cs) are used to identify those isolates which now fail to grow (conditional mutants). As above close monitoring of the changes upon growth at the non-permissive temperature yields information on the function of the mutated gene. Complementation of conditional lethal mutation by library from target organism and sequencing of complementing gene allows matching with library sequences.

Each of these techniques may have advantages or disadvantages depending on the particular application. The skilled artisan would choose the approach that is the most relevant with the particular end use in mind. For example, some genes might be recognised as essential for infection but in reality are only necessary for the initiation of infection and so their products would represent relatively unattractive targets for antibacterials developed to cure established and chronic infections.

#### 6) RT-PCR

Bacterial messenger RNA, preferably that of *S. epidermidis*, is isolated from bacterial infected tissue e.g. 48 hour murine lung infections, and the amount of each mRNA species assessed by reverse transcription of the RNA sample primed with random hexanucleotides followed by PCR with gene specific primer pairs. The determination of the presence and amount of a particular mRNA species by quantification of the resultant PCR product provides information on the bacterial genes which are transcribed in the infected tissue. Analysis of gene transcription can be carried out at different times of infection to gain a detailed knowledge of gene regulation in bacterial pathogenesis allowing for a clearer understanding of which gene products represent targets for screens for novel antibacterials. Because of the gene specific nature of the PCR primers employed it should be understood that the bacterial mRNA preparation need not be free of mammalian RNA. This allows the investigator to carry out a simple and quick RNA preparation from infected tissue to obtain bacterial mRNA species which are very short lived in the bacterium (in the order of 2 minute half-lives). Optimally the bacterial mRNA is prepared from infected murine lung tissue by mechanical disruption in the presence of TRIzol (GIBCO-BRL) for very short periods of time, subsequent processing according to the manufacturers of TRIzol reagent and

DNAase treatment to remove contaminating DNA. Preferably the process is optimized by finding those conditions which give a maximum amount of bacterial 16S ribosomal RNA, preferably that of *S. epidermidis*, as detected by probing Northernblots with a suitably labeled sequence specific oligonucleotide probe.

- 5 Typically a 5' dye labelled primer is used in each PCR primer pair in a PCR reaction which is terminated optimally between 8 and 25 cycles. The PCR products are separated on 6% polyacrylamide gels with detection and quantification using GeneScanner (manufactured by ABI).

- 10 Use of the of these technologies when applied to the sequences of the present invention enables identification of bacterial proteins expressed during infection, inhibitors of which would have utility in anti-bacterial therapy.

### Polynucleotides

- The present invention relates to novel polynucleotides and novel  
15 polypeptides of *S. epidermidis*, among other things, as described below. The invention particularly relates to the nucleotide sequences set forth in the Sequence Listing SEQ ID NOs: 1-3334, typically as odd numbered ID numbers, and the corresponding deduced amino acid sequences also set forth in the Sequence Listing SEQ ID NOs:1-3334, typically as even numbered ID numbers.  
20 SEQ ID NOs 1-3334 refer to open reading frames (ORFs). The invention also relates to consensus polynucleotide sequences from which the ORFs were extracted. These genomic sequences include the ORFs, intergenic regions and ribosomal RNA genes. Such genomic polynucleotides are set forth as SEQ ID Nos 3335-4464. It will be noted that minor errors in sequencing can occur which  
25 do not depart from the spirit of the invention; *S. epidermidis* polynucleotides and polypeptides having any corrected sequences are thus encompassed by this invention.

- Using the information provided herein and known, standard methods, such as those for cloning and sequencing and those for synthesizing polynucleotides  
30 and polypeptides (see, e.g., Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)), one can generate numerous unique fragments, both longer and shorter than the polynucleotides and polypeptides set forth in the Sequence Listing, of the *S. epidermidis* genome and the *S. epidermidis* coding regions, which are  
35 encompassed by the present invention. To be unique, a fragment must be of sufficient size to distinguish it from other known nucleic acid sequences, most

readily determined by comparing any selected *S. epidermidis* fragment to the nucleotide sequences in computer databases such as GenBank. Such comparative searches are standard in the art. Many unique fragments will be *S. epidermidis* – specific. Typically, a unique fragment useful as a primer or probe will be at least about 20 to about 25 nucleotides in length, depending upon the specific nucleotide content of the sequence. Additionally, fragments can be, for example, at least about 30, 40, 50, 60, 75, 80, 90, 100, 150, 200, 250, 300, 400, 500 or more nucleotides in length. The nucleic acid fragment can be single, double or triple stranded, depending upon the purpose for which it is intended.

Additionally, as discussed above and below, modifications can be made to the *S. epidermidis* polynucleotides and polypeptides that are encompassed by the present invention. For example, nucleotide substitutions can be made which do not affect the polypeptide encoded by the nucleic acid, and thus any polynucleotide which encodes the polypeptides of this invention is within the present invention. Additionally, certain amino acid substitutions (and corresponding nucleotide substitutions to encode them) can be made which are known in the art to be neutral (Robinson W.E. Jr. and Mitchell, W.m., *AIDS* 4: S141-S162 (1990). Such variations may arise naturally as allelic variations (*e.g.*, due to genetic polymorphism) or may be produced by human intervention (*e.g.*, by mutagenesis of cloned DNA sequences), such as induced point, deletion, insertion and substitution mutations. Minor changes in amino acid sequence are generally preferred, such as conservative amino acid replacements, small internal deletions or insertions, and additions or deletions at the ends of the molecules. Substitutions may be designed based on, for example, the model of Dayhoff, *et al.* (in *Atlas of Protein Sequence and Structure* 1978, Nat'l Biomed. Res. Found., Washington D.C.). These modifications can result in changes in the amino acid sequence, provide silent mutations, modify a restriction site, or provide other specific mutations. Likewise, such amino acid changes result in a different nucleic acid encoding the polypeptides and proteins. Thus, alternative polynucleotides, which are within the parameters of the present invention, are contemplated by such modifications.

Furthermore, the polynucleotide sequences set forth as SEQ ID Nos: 1-3334 in the Sequence Listing are open reading frames (ORFs), *i.e.*, coding regions of *S. epidermidis*. The polypeptide encoded by each open reading frame can be deduced, and the molecular weight of the polypeptide thus calculated using amino acid residue molecular weight values well known in the art. Any

selected coding region can be functionally linked, using standard techniques such as standard subcloning techniques, to any desired regulatory sequence, whether a *S. epidermidis* regulatory sequence or a heterologous regulatory sequence, or to a heterologous coding sequence to create a fusion protein, as further described herein.

Polynucleotides of the present invention may be in the form of RNA, such as mRNA or cRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be triple-stranded, double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The coding sequence which encodes a *S. epidermidis* polypeptide of this invention may be identical to the coding sequence of a polynucleotide set forth in the sequence listing. It also may be a polynucleotide with a different sequence which, as a result of the redundancy (degeneracy) of the genetic code, encodes a *S. epidermidis* polypeptide set forth in the sequence listing.

Polynucleotides of the present invention which encode a *S. epidermidis* polypeptide set forth in the sequence listing may include, but are not limited to, the coding sequence for a mature polypeptide, by itself; the coding sequence for a mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of a mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription (including termination signals, for example), ribosome binding, mRNA stability elements, and additional coding sequence which encode additional amino acids, such as those which provide additional functionalities. Thus, for instance, a polypeptide may be fused to a marker sequence, such as a peptide, which facilitates purification of the fused polypeptide. In certain embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in the pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The HA tag may also be used to create fusion proteins and corresponds to an epitope derived of influenza hemagglutinin protein, which has been described by

Wilson *et al.*, *Cell* 37: 767 (1984), for instance. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated genetic elements.

In accordance with the foregoing, the term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides which include a sequence encoding a polypeptide of the present invention, particularly a polypeptide having a *S. epidermidis* amino acid sequence set forth in the Sequence Listing. The term encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by integrated phage or insertion sequence or editing) together with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the herein above described polynucleotides which encode for fragments, analogs and derivatives of the polypeptide having a deduced *S. epidermidis* amino acid sequence set forth in the Sequence Listing. A variant of the polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the polynucleotide may be made by mutagenesis techniques, including those applied to polynucleotides, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned polynucleotides by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Preferred are polynucleotides encoding a variant, analog, derivative or fragment, or a variant, analogue or derivative of a fragment, which have a *S. epidermidis* sequence as set forth in the Sequence Listing, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid(s) is substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the *S. epidermidis* polypeptides set forth in the Sequence Listing. Also especially preferred in this regard are conservative substitutions.

Further preferred embodiments of the invention are polynucleotides that are at least 70% identical over their entire length to a polynucleotide encoding a polypeptide having an amino acid sequence set forth in the Sequence Listing, and polynucleotides which are complementary to such polynucleotides. Alternatively,

most highly preferred are polynucleotides that comprise a region that is at least 80% or at least 85% identical over their entire length to a polynucleotide encoding a *S. epidermidis* polypeptide set forth in the Sequence Listing, including complementary polynucleotides. In this regard, polynucleotides at least 90%,  
5 91%, 92%, 93%, 94%, 95%, or 96% identical over their entire length to the same are particularly preferred, and among these particularly preferred polypeptides, those with at least 95% are especially preferred. . Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these, those with at least 98% and at least 99% are particularly highly preferred, with at least 99%  
10 or 99.5% being the more preferred.

Preferred embodiments in this respect, moreover, are polynucleotides which encode polypeptides which retain substantially the same biological function or activity as the mature polypeptide encoded by the DNA set forth in the Sequence Listing.

15 The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. Stringent conditions are typically selective conditions. As herein used, the term "stringent conditions" means hybridization will  
20 occur only if there is at least 95% and preferably at least 97% identity between the sequences. For a specific sequence, stringent conditions can be determined empirically according to the nucleotide content, as is known in the art. For example, a typical example of stringent conditions is hybridization of a 48mer having 55% GC content at 42°C in 50% formamide and 750 mM NaCl followed by washing at 55°C in  
25 15 mM NaCl and 0.1% SDS.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to  
30 isolate cDNA and genomic clones of other genes that have a high sequence similarity to the polynucleotides of the present invention. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 20, at least 25 or at least 30 bases, and may have at least 50 bases. Particularly preferred probes will have at least 30 bases, and will have 50 bases or less, such as 30, 35, 40, 45, or  
35 50 bases.



For example, the coding region of the polynucleotide of the present invention may be isolated by screening using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the present invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine to which members of the library the probe hybridizes.

The polynucleotides and polypeptides of the present invention may be employed as reagents and materials for development of treatments of and diagnostics for disease, particularly human disease, as further discussed herein relating to polynucleotide assays, *inter alia*.

The polynucleotides of the present invention that are oligonucleotides can be used in the processes herein as described, but preferably for PCR, to determine whether or not the *S. epidermidis* genes identified herein in whole or in part are present and/or transcribed in infected tissue such as blood. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The polynucleotides may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids; or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

The present invention additionally contemplates polynucleotides functionally encoding fusion polypeptides wherein the fusion polypeptide comprises a fragment of a *S. epidermidis* polypeptide and one or more polypeptide(s) derived from another *S. epidermidis* polypeptide or from another organism or a synthetic polyamino acid sequence. Such polynucleotides may or may not encode amino acid sequences to facilitate cleavage of the *S. epidermidis* polypeptide from the other polypeptide(s) under appropriate conditions.

In sum, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

### Polypeptides

The present invention further relates to peptides, polypeptides and proteins (collectively referred to as "polypeptides") of *S. epidermidis*. The amino acid sequence of these polypeptides is set forth in the Sequence Listing.

The invention also relates to fragments, analogs and derivatives of these polypeptides. The terms "fragment," "derivative" and "analog" when referring to a polypeptide whose amino acid sequence is set forth in the Sequence Listing, means a polypeptide which retains essentially the same biological function or activity as such polypeptide. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The fragment, derivative or analog of the polypeptide of the present invention may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Among the particularly preferred embodiments of the invention in this regard are polypeptides set forth in the Sequence Listing, variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of the fragments. Additionally, fusion polypeptides comprising such polypeptides, variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of the fragments, in addition to a heterologous polypeptide, are contemplated by the

present invention. Such fusion polypeptides and proteins, as well as polynucleotides encoding them, can readily be made using standard techniques, including standard recombinant techniques for producing and expressing a recombinant polynucleic acid encoding a fusion protein.

5        Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl  
10 residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

Further particularly preferred in this regard are variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, having the  
15 amino acid sequence of any polypeptide set forth in the Sequence Listing, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the polypeptide of the present invention. Also especially preferred in this regard  
20 are conservative substitutions. Most highly preferred are polypeptides having an amino acid sequence set forth in the Sequence Listing without substitutions.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include any polypeptide set forth in  
25 the Sequence Listing (in particular a mature polypeptide) as well as polypeptides which have at least 70% identity to a polypeptide set forth in the Sequence Listing, preferably at least 80% or 85% identity to a polypeptide set forth in the Sequence Listing, and more preferably at least 90% similarity (more preferably at least 90% identity) to a polypeptide set forth in the Sequence Listing and still more preferably at  
30 least 95%, 96%, 97%, 98%, 99%, or 99.5% similarity (still more preferably at least 95%, 96%, 97%, 98%, 99%, or 99.5% identity) to a polypeptide set forth in the Sequence Listing and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids, such as 30, 35, 40, 45 or 50 amino acids.

35        Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide

synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

5

### **Fragments**

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of the polypeptide having the amino acid sequence set forth in the Sequence Listing, and fragments of variants and derivatives of the polypeptides set forth in the Sequence Listing.

10

In this regard a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned *S. epidermidis* polypeptides and variants or derivatives thereof.

15

Such fragments may be "free-standing," *i.e.*, not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the presently discussed fragments most preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of a polypeptide of the present invention comprised within a precursor polypeptide designed for expression in a host and having heterologous pre and pro-polypeptide regions fused to the amino terminus of the fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one aspect of the meaning intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from a polypeptide of the present invention.

25

Representative examples of polypeptide fragments of the invention, include, for example, in any selected polypeptide, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, and 101-200, 201-300, or, at the COOH-terminal end, the C-terminal 20 amino acids, the C-terminal 30 amino acids, the C-terminal 40 amino acids, the C-terminal 50 amino acids, and any combination of these.

30

fragments, such as fragment from about amino acid number 1-40, 1-60, 21-60, 41-80, 61-100, and the like.

In this context "about" herein includes the particularly recited ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

35

Preferred fragments of the invention include, for example, truncation polypeptides including polypeptides having an amino acid sequence set forth in the

Sequence Listing, or of variants or derivatives thereof, except for deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Fragments having the size ranges set out above also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally. Degradation forms of the polypeptides of the invention in a host cell are also preferred.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of the polypeptide of the present invention. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions of the polypeptide of the present invention, and combinations of such fragments.

Preferred regions are those that mediate activities of the polypeptide of the present invention. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of the polypeptide of the present invention, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Particularly preferred are fragments comprising receptors or domains of enzymes that confer a function essential for viability of *S.epidermidis* or the ability to cause disease in humans. Further preferred polypeptide fragments are those that comprise or contain antigenic or immunogenic determinants in an animal, especially in a human.

It will be appreciated that the invention also relates to, among others, polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides, such as PCR primers, for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspond to the preferred fragments, as discussed above.

**Vectors, host cells, expression**

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. Introduction of a polynucleotides into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Polynucleotide constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, a single or double-stranded RNA or DNA viral vector. Plasmids generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially available, publicly available, or can be constructed from available plasmids by routine application of well known, published procedures, given the teachings herein. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art.

Preferred among vectors, in certain respects, are those for expression of polynucleotides and polypeptides of the present invention. Generally, such vectors comprise *cis*-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate *trans*-acting factors either  
5 are supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific.  
10 Particularly preferred among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of  
15 skill in the art.

A great variety of expression vectors can be used to express a polypeptide of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements,  
20 from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with  
25 this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques, such as, for example, those set forth in  
30 Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representatives of such promoters include, but are not  
35 limited to, the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs.

In general, expression constructs will contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon  
5 appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions will operate by controlling transcription, such as transcription factors, repressor binding sites and termination, among others.

10 Vectors for propagation and expression generally will include selectable markers and amplification regions, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

Representative examples of appropriate hosts include bacterial cells, such as  
15 streptococci, staphylococci, *E. coli*, streptomyces and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

The following vectors, which are commercially available, are provided by way  
20 of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia, and pBR322 (ATCC 37017). Among preferred eukaryotic vectors are pWLNEO, pSV2CAT,  
25 pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example,  
30 introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("CAT") transcription unit, downstream of  
35 restriction site or sites for introducing a candidate promoter fragment; *i.e.*, a fragment that may contain a promoter. As is well known, introduction into the vector of a



promoter-containing fragment at the restriction site upstream of the *cat* gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available, such as pKK232-8 and pCM7. Promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

Among known prokaryotic promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* *lacI* and *lacZ* and promoters, the T3 and T7 promoters, the *gpt* promoter, the lambda PR, PL promoters and the *trp* promoter.

Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Recombinant expression vectors will include, for example, origins of replication, a promoter preferably derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

Polynucleotides of the invention, encoding the heterologous structural sequence of a polypeptide of the invention generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates translation of the polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and the initiation codon. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal in constructs for use in eukaryotic hosts. Transcription termination signal appropriately disposed at the 3' end of the transcribed region may also be included in the polynucleotide construct.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide.

These signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N- or C-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, regions also may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability or to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to solubilize or purify polypeptides. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another protein or part thereof. In drug discovery, for example, proteins have been fused with antibody Fc portions for the purpose of high-throughput screening assays to identify antagonists. See, D. Bennett *et al.*, *Journal of Molecular Recognition*, Vol. 8 52-58 (1995) and K. Johanson *et al.*, *The Journal of Biological Chemistry*, Vol. 270, No. 16, pp 9459-9471 (1995).

Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents; such methods are well known to those skilled in the art.

Mammalian expression vectors may comprise expression sequences, such as an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation regions, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences that are useful or necessary for expression.

The polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high

performance liquid chromatography is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

#### **Polynucleotide assays**

5           This invention is also related to the use of the polynucleotides of the present invention to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of complementary nucleotides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of a disease. Eukaryotes (herein also "individual(s)"), particularly  
10   mammals, and especially humans, infected with *S. epidermidis* may be detected at the DNA level by a variety of techniques. By selecting regions of nucleic acids that vary among strains of *S. epidermidis*, preferred candidates for distinguishing a specific strain of *S. epidermidis* can be obtained. Furthermore, by selecting regions of nucleic acids that vary between *S. epidermidis* and other organisms, preferred  
15   candidates for distinguishing a *S. epidermidis* from other organisms can be obtained. Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as bone, blood, muscle, cartilage, and skin. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki et al., *Nature*, 324: 163-166 (1986) prior to analysis. RNA or cDNA may also be used  
20   in the same ways. As an example, PCR primers complementary to the nucleic acid forming part of the polynucleotide of the present invention can be used to identify and analyze for its presence and/or expression. Using PCR, characterization of the strain of *S. epidermidis* present in a mammal, and especially a human, may be made by an analysis of the genotype of the prokaryote gene. For example, deletions and  
25   insertions can be detected by a change in size of the amplified product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing amplified DNA to radiolabeled RNA or alternatively, radiolabeled antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in  
30   melting temperatures.

          Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or  
35   another amplification method. For example, a sequencing primer can be used with double-stranded PCR product or a single-stranded template molecule generated by

a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic characterization based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers et al., *Science*, 230: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton et al., *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, e.g., restriction fragment length polymorphisms (RFLP) and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by *in situ* analysis.

Cells carrying mutations or polymorphisms in the gene of the present invention may also be detected at the DNA level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA or cDNA may also be used for the same purpose, PCR or RT-PCR. As an example, PCR primers complementary to the nucleic acid encoding the polypeptide of the present invention can be used to identify and analyze mutations. The primers may be used to amplify the gene isolated from the individual such that the gene may then be subject to various techniques for elucidation of the DNA sequence. In this way, mutations in the DNA sequence may be diagnosed.

The invention provides a process for diagnosing disease, arising from infection with *S. epidermidis*, comprising determining from a sample isolated or derived from an individual an increased level of expression of a polynucleotide having the sequence of a polynucleotide set forth in the Sequence Listing. Increased expression of polynucleotide can be measured using any one of the

methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

### **Polypeptide assays**

5       The present invention also relates to diagnostic assays such as quantitative and diagnostic assays for detecting levels of the polypeptide of the present invention in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression of the polypeptide compared to normal control tissue samples may be  
10       used to detect the presence of an infection, for example, and to identify the infecting organism. Assay techniques that can be used to determine levels of a polypeptide, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these, ELISAs frequently are preferred. An  
15       ELISA assay initially comprises preparing an antibody specific to the polypeptide, preferably a monoclonal antibody. In addition, a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached to a detectable reagent such as radioactive, fluorescent or enzymatic reagent, such as horseradish peroxidase enzyme.

### **Antibodies**

20       The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. The present invention includes, for example, monoclonal and polyclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the  
25       product of an Fab expression library.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptide itself. In this manner, even a  
30       sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique known in the art which provides antibodies produced by continuous cell line cultures can be used.  
35       Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et*

al., pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985); U.S. Patent No. 5,545,403; U.S. Patent No. 5,545,405; U.S. Patent No. 5,654,403; U.S. Patent No. 5,792,838; U.S. Patent No. 5,316,938; U.S. Patent No. 5,633,162; U.S. Patent No. 5,644,036; U.S. Patent No. 5,858,725.

5           Techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

10           Alternatively, phage display technology could be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-Fbp or from naive libraries (McCafferty, J. et al., (1990), *Nature* **348**, 552-554; Marks, J. et al., (1992) *Biotechnology* **10**, 779-783). The affinity of these  
15           antibodies can also be improved by chain shuffling (Clackson, T. et al., (1991) *Nature* **352**, 624-628).

          If two antigen binding domains are present, each domain may be directed against a different epitope - termed 'bispecific' antibodies.

          The above-described antibodies may be employed to isolate or to identify  
20           clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

          Thus, among others, antibodies against the polypeptide of the present invention may be employed to inhibit and/or treat infections, particularly bacterial  
25           infections and especially infections arising from *S. epidermidis*.

          Polypeptide derivatives include antigenically, epitopically or immunologically equivalent derivatives which form a particular aspect of this invention. The term "antigenically equivalent derivative" as used herein encompasses a polypeptide or its equivalent which will be specifically recognized  
30           by certain antibodies which, when raised to the protein or polypeptide according to the present invention, interfere with the immediate physical interaction between pathogen and mammalian host. The term "immunologically equivalent derivative" as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to  
35           interfere with the immediate physical interaction between pathogen and mammalian host.

The polypeptide, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof can be used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by  
5 conjugation, with an immunogenic carrier protein, for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively, a multiple antigenic peptide comprising multiple copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof, may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a  
10 carrier.

Preferably the antibody or derivative thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," wherein the complementarity determining region(s) of the hybridoma-derived antibody has been transplanted  
15 into a human monoclonal antibody, for example as described in Jones, P. et al. (1986), *Nature* **321**, 522-525 or Tempest et al., (1991) *Biotechnology* **9**, 266-273.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscle (Wolff et al., *Hum Mol Genet* 1992, 1:363, Manthorpe et al.,  
20 *Hum. Gene Ther.* 1993:4, 419), delivery of DNA complexed with specific protein carriers (Wu et al., *J Biol Chem* 1989:264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS*, 1986:83,9551), encapsulation of DNA in various forms of liposomes (Kaneda et al., *Science* 1989:243,375), particle bombardment (Tang et al., *Nature* 1992, 356:152, Eisenbraun et al., *DNA*  
25 *Cell Biol* 1993, 12:791) and *in vivo* infection using cloned retroviral vectors (Seeger et al., *PNAS* 1984:81,5849).

#### **Binding molecules and assays**

This invention also provides a method for identification of molecules, such as binding molecules, that bind to the polypeptide of the present invention. Genes  
30 encoding proteins that bind to the polypeptide can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan et al., *Current Protocols in Immunology* 1(2): Chapter 5 (1991). Also, a labeled ligand can be photoaffinity linked to a cell extract. Polypeptides of  
35 the invention also can be used to assess the binding capacity of a binding molecule, in cells or in cell-free preparations.

Polypeptides of the invention may also be used to assess the binding or small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics.

5       The invention further provides a complex of a polypeptide and a binding molecule which comprises a polypeptide as described herein and a binding molecule capable of modulating the activity of the polypeptide. A complex of this kind can arise *in vivo* upon administration to a patient of a binding molecule as described herein.

#### 10       **Antagonists and agonists - assays and molecules**

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the function of polypeptides or polynucleotides of the present invention, such as its interaction with a binding molecule. The method of screening may involve high-throughput.

15       For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, may be prepared from a cell that expresses a molecule that binds to the polypeptide of the present invention. The preparation is incubated with labeled polypeptide in the absence or the presence of a candidate molecule  
20       which may be an agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, *i.e.*, without inducing the functional effects of the polypeptide, are most likely to be good antagonists. Molecules that bind well and elicit functional effects that are the same as or closely related to the polypeptide are  
25       good agonists.

The functional effects of potential agonists and antagonists may be measured, for instance, by determining activity of a reporter system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of the polypeptide of the present invention or  
30       molecules that elicit the same effects as the polypeptide. Reporter systems that may be useful in this regard include but are not limited to colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in the functional activity of the polypeptide, and binding assays known in the art.

Another example of an assay for antagonists is a competitive assay that  
35       combines the polypeptide of the present invention and a potential antagonist with membrane-bound binding molecules, recombinant binding molecules, natural



substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. The polypeptide can be labeled such as by radioactivity or a colorimetric compound, such that the number of polypeptide molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds to the same sites on a binding molecule without inducing functional activity of the polypeptide of the invention.

Potential antagonists include a small molecule which binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules.

Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), for a description of these molecules).

Preferred potential antagonists include derivatives of the polypeptide of the invention.

In a particular aspect, the invention provides the use of the polypeptide, polynucleotide or inhibitor of the invention to interfere with the initial physical interaction between a pathogen and mammalian host responsible for sequelae of infection. In particular the molecules of the invention may be used: i) in the prevention of adhesion of *S. epidermidis* to mammalian extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; ii) to block protein mediated mammalian cell invasion by, for example, initiating phosphorylation of mammalian tyrosine kinases (Rosenshine *et al.*, *Infect. Immun.* 60:2211 (1992)); iii) to block bacterial adhesion between mammalian extracellular matrix proteins and bacterial proteins which mediate tissue damage; iv) to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

Each of the DNA coding sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein

upon expression can be used as a target for the screening of antibacterial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The antagonists and agonists may be employed, for instance, to inhibit diseases arising from infection with *Staphylococcus*, especially *S. epidermidis*, such as sepsis and endocarditis.

#### Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, which comprises inoculating the individual with the polypeptide of the invention, or a fragment or variant thereof, adequate to produce antibody to protect said individual from infection, particularly bacterial infection and most particularly *Staphylococcus* infections. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises, through gene therapy or otherwise, delivering a nucleic acid functionally encoding the polypeptide, or a fragment or a variant thereof, for expressing the polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response to produce antibodies or a cell mediated T cell response, either cytokine-producing T cells or cytotoxic T cells, to protect said individual from disease, whether that disease is already established within the individual or not. One way of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise.

A further aspect of the invention relates to an immunological composition which, when introduced into a host capable of having induced within it an immunological response, induces an immunological response in such host, wherein the composition comprises recombinant DNA which codes for and expresses an antigen of the polypeptide of the present invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity or cellular immunity such as that arising from CTL or CD4+ T cells.

The polypeptide of the invention or a fragment thereof may be fused with co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a fused protein which will have immunogenic and protective properties. This fused recombinant protein preferably further comprises

an antigenic co-protein, such as Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilise the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof which have been shown to encode non-variable regions of bacterial cell surface proteins in DNA constructs used in such genetic immunization experiments in animal models of infection with *S. epidermidis*. Such fragments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. This approach can allow for the subsequent preparation of monoclonal antibodies of particular value from the requisite organ of the animal successfully resisting or clearing infection for the development of prophylactic agents or therapeutic treatments of *S. epidermidis* infection in mammals, particularly humans.

The polypeptide may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue. Examples of tissue damage include wounds in skin or connective tissue caused e.g. by mechanical, chemical or thermal damage or by implantation of indwelling devices, or wounds in the mucous membranes, such as the mouth, mammary glands, urethra or vagina.

The present invention also includes a vaccine formulation which comprises the immunogenic recombinant protein together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include

suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials, and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include  
5    adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in-water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

While the invention has been described with reference to certain  
10    polypeptides, it is to be understood that this covers fragments of the naturally occurring protein and similar proteins with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant protein.

### Compositions

15    The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the agonists or antagonists. Thus, the polypeptides of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions  
20    comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

### 25    Kits

The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. The ingredient(s) can be present in a useful amount, dosage, formulation or combination. Associated with such  
30    container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

### Administration

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

5       The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

10       The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. In general, the compositions are administered in an amount of active agent of at least about 10  $\mu\text{g/kg}$  body weight. In most cases they will be administered in one or more doses in an amount not in excess of about 8 mg/kg body weight per day. Preferably, in most cases, dose is from about 10  $\mu\text{g/kg}$  to about 1 mg/kg body weight, daily. For  
15       administration particularly to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg and typically around 1 mg/kg. For example, a dose may be 1 mg/kg daily. It will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity,  
20       route of administration, complicating conditions and the like. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such  
25       are within the scope of this invention.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

30       Alternatively the composition may be formulated for topical application, for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for  
35       example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such

carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

The pharmaceutical composition may be administered in conjunction with an in-dwelling device. In-dwelling devices include surgical implants, prosthetic devices and catheters, i.e., devices that are introduced to the body of an individual and remain in position for an extended time. Such devices include, for example, artificial joints, heart valves, pacemakers, vascular grafts, vascular catheters, cerebrospinal fluid shunts, urinary catheters, continuous ambulatory peritoneal dialysis (CAPD) catheters, etc.

The composition of the invention may be administered by injection to achieve a systemic effect against relevant bacteria shortly before insertion of an in-dwelling device. Treatment may be continued after surgery during the in-body time of the device. In addition, the composition could also be used to broaden perioperative cover for any surgical technique to prevent *Staphylococcus* wound infections.

Many orthopaedic surgeons consider that humans with prosthetic joints should be considered for antibiotic prophylaxis before dental treatment that could produce a bacteremia. Late deep infection is a serious complication sometimes leading to loss of the prosthetic joint and is accompanied by significant morbidity and mortality. It may therefore be possible to extend the use of the active agent as a replacement for prophylactic antibiotics in this situation.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of bacteria to matrix proteins exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion. The active agent will preferably be present at a concentration of  $1\mu\text{g/ml}$  to  $10\text{mg/ml}$  for bathing of wounds or indwelling devices.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is  $0.5\text{-}5\mu\text{g/kg}$  of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks.

With the indicated dose range, no adverse toxicological effects should be observed with the compounds of the invention which would preclude their administration to suitable individuals.

The antibodies described above may also be used as diagnostic reagents to detect the presence of bacteria containing the protein.

In order to facilitate understanding of the following example certain frequently occurring methods and/or terms are explained in the foregoing glossary.

5       The present invention is further described by the following examples. While illustrating certain specific aspects of the invention, the examples do not portray the limitations or circumscribe the scope of the disclosed invention.

All examples were carried out using routine molecular biology techniques as generally described in standard laboratory manuals, such as Sambrook et al.,  
10    *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

### EXAMPLES

15       A small insert plasmid library was generated in the minimal sequencing vector pOT2a (O. Hubbard, C. Martin, and M. Palazzolo, unpublished). pOT2a vector was prepared by BstXI digestion of the parent plasmid pOT2a-sacB followed by preparative agarose gel electrophoresis to separate the 1.6 kb vector fragment from a *B. subtilis* sacB gene fragment. To prepare inserts for library  
20    construction *S. epidermidis* SR1 strain genomic DNA was sonicated and the resulting random fragments were end-repaired with klenow and T4 polymerase and phosphorylated with T4 polynucleotide kinase. Oligos (5'-CTCTAAAG-3', 5'-CTTTAGAGCACA-3') (SEQ ID NO.:4465) to create BstXI adaptors were annealed and ligated to the blunt-ended fragments. The configuration of the BstXI  
25    sites in pOT2a and the sequence of the adaptors allowed a ligation strategy that minimized the recovery of clones without insert (Seed, 1987). DNA samples were electrophoresed on a low-melting-temperature agarose gel and fragments of 3000-4000 bp were isolated and purified. The linearized vector and random DNA fragments were ligated overnight using T4 DNA ligase at 16C and transformed  
30    into DH10B competent cells (Life Technologies Inc., Gaithersburg, MD) by electroporation. Transformed bacteria were selected on LB agar plates containing 5% sucrose and 12.5 ug/ml chloramphenicol. Sequencing templates were isolated from single colonies and purified using R.E.A.L. Prep 96 Plasmid Kit (QIAGEN, Chatsworth, CA). Seq01 primer (5'-CACTATAGAACTCGAGCAGCTG-  
35    3') (SEQ ID NO.:4466) and seq02 primer (5'-CGACTCACTATAGGGAGACCG-3')

(SEQ ID NO.:4467) were used to generate end-sequence using ABI Prism BigDye Terminators (PE Applied Biosystems, Foster City, CA).

Constructs from the pOT2a library were transformed into POX38 bacteria and selected on LB agar plates containing 12.5 ug/ml chloramphenicol. A single colony from each construct was used to inoculate an overnight culture. These POX38 cultures were mated with a culture of the F-bearing, kanamycin resistant JGM strain by combining the two strains and shaking for 3 hours at 37C without antibiotics. Each successful mating event resulted in the random insertion of a single gamma-delta transposon into the pOT2a construct. This collection of transpositions was captured in the JGM cells by selection of the mated cultures on LB agar plates containing 12.5 ug/ml chloramphenicol and 25 ug/ml kanamycin. A transposon library was created for each of the original pOT2a library constructs by picking 96 individual colonies. A set of two PCR reactions was performed on each of the 96 library members to determine the position of the transposon integration. PM001 primer (5'- CGTTAGAACGCGGCTACA-3') (SEQ ID NO.:4468) and NGDIR primer (5'-GTTCCATTGGCCCTCAAAC-3') (SEQ ID NO.:4469) were used to determine the integration site distance from the left side of the vector and PM002 primer (5'-GCCGATTCATTAATGCAGGT-3') (SEQ ID NO.:4470) and NGDIR primer were used to confirm the integration position by measuring the distance from the right side of the vector. PCR products were electrophoresed in 1X TBE on 1.4% agarose gels. After gel analysis, a subset of transposon clones was selected for sequencing based upon obtaining an integration site about every 300 bp along the full length of the pOT2a insert. Sequencing templates were purified using R.E.A.L. Prep 96 Plasmid Kit (QIAGEN, Chatsworth, CA). M21 primer (5'-GTAAAACGACGGCCAGT-3') (SEQ ID NO.:4471) and rev primer (5'-CAGGAAACAGCTATGAC-3') (SEQ ID NO.:4472) were used to generate internal sequence using ABI Prism BigDye Terminators (PE Applied Biosystems, Foster City, CA).

The sequences, including ORFs (nucleic acid sequences within SEQ ID NOs 1-3334 ) and non-ORFs (SEQ ID NOs 3335-4464 ) are set forth in the Sequence Listing. The non-ORF regions may be particularly useful as diagnostic sequences. The ribosomal RNA genes may also be useful to distinguish between species. Also, intergenic regions generally may be useful diagnostics to establish genus and species of an unidentified microbe, as there may be less selective pressure to maintain fidelity of the sequences in these intergenic regions as compared to intragenic regions.



About 26 different isolates of *S. epidermidis* have been submitted to ATCC listed in their on-line catalog, listed below:

- 1: ATCC Number: 146 Organism: *Staphylococcus epidermidis*
- 2: ATCC Number: 33501 Organism: *Staphylococcus epidermidis*
- 5 3: ATCC Number: 49741 Organism: *Staphylococcus epidermidis*
- 4: ATCC Number: 51625 Organism: *Staphylococcus epidermidis*
- 5: ATCC Number: 29997 Organism: *Staphylococcus epidermidis*
- 6: ATCC Number: 19654 Organism: *Staphylococcus epidermidis*
- 7: ATCC Number: 14389 Organism: *Staphylococcus sp. deposit*
- 10 8: ATCC Number: 14852 Organism: *Staphylococcus epidermidis*
- 9: ATCC Number: 49134 Organism: *Staphylococcus epidermidis*
- 10: ATCC Number: 13518 Organism: *Staphylococcus epidermidis*
- 11: ATCC Number: 9491 Organism: *Staphylococcus epidermidis*
- 12: ATCC Number: 35547 Organism: *Staphylococcus epidermidis*
- 15 13: ATCC Number: 35984 Organism: *Staphylococcus epidermidis*
- 14: ATCC Number: 35983 Organism: *Staphylococcus epidermidis*
- 15: ATCC Number: 700296 Organism: *Staphylococcus epidermidis*
- 16: ATCC Number: 49461 Organism: *Staphylococcus epidermidis*
- 17: ATCC Number: 29641 Organism: *Staphylococcus epidermidis*
- 20 18: ATCC Number: 29887 Organism: *Staphylococcus epidermidis*
- 19: ATCC Number: 29886 Organism: *Staphylococcus epidermidis*
- 20: ATCC Number: 55133 Organism: *Staphylococcus epidermidis*
- 21: ATCC Number: 27626 Organism: *Staphylococcus sp. deposit*
- 22: ATCC Number: 31874 Organism: *Staphylococcus epidermid*
- 25 23: ATCC Number: 14990 Organism: *Staphylococcus epidermid*
- 24: ATCC Number: 155 Organism: *Staphylococcus sp. deposit*
- 25: ATCC Number: 155-U Organism: *Staphylococcus sp. depos*
- 26: ATCC Number: 12228 Organism: *Staphylococcus epidermid*

30 Throughout this application, various publications are referenced. These publications are hereby incorporated by reference in their entirety.

While the invention has been described with respect to certain specific embodiments, it will be appreciated that many modifications and changes may be made by those skilled in the art without departing from the spirit of the invention.

35 It is intended, therefore, by the appended claims, to cover all such modification and changes as fall within the true spirit and scope of the invention.

**What is claimed is:**

1. An isolated polynucleotide comprising a member selected from the group consisting of:
  - (a) a polynucleotide encoding a polypeptide having at least a 70% identity to a polypeptide set forth in the Sequence Listing;
  - (b) a polynucleotide which is complementary to the polynucleotide of (a); and
  - (c) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a) or (b).
2. The polynucleotide of Claim 1 wherein the polypeptide has at least 80% identity to the polypeptide set forth in the Sequence Listing.
3. The polypeptide of Claim 2 wherein the polypeptide has at least 90% identity to the polypeptide set forth in the Sequence Listing.
4. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.
5. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.
6. The polynucleotide of Claim 4 wherein the polynucleotide has at least 80% identity to a polynucleotide set forth in the Sequence Listing.
7. The polynucleotide of Claim 6 wherein the polynucleotide has at least 90% identity to a polynucleotide set forth in the Sequence Listing.
8. The polynucleotide of Claim 4 comprising a polynucleotide set forth in the Sequence Listing.
9. The polynucleotide of Claim 4 comprising a polynucleotide set forth in the Sequence Listing as any of an odd-numbered SEQ ID Nos 1-3334.
10. An isolated polynucleotide comprising a member selected from the group consisting of:
  - (a) a polynucleotide having at least a 70% identity to a polynucleotide contained in any of ATCC Deposit Nos. 146; 33501; 49741; 51625; 29997; 19654 ; 14389 ; 14852 ; 9134 ; 13518; 9491; 35547 ; 35984; 35983; 700296; 49461; 29641; 29887 ; 29886; 55133; 27626; 31874; 14990; 155; 155-U; 12228 and substantially encoding the polypeptide comprising amino acids 1 to 416 of SEQ ID NO:2;
  - (b) a polynucleotide complementary to the polynucleotide of (a); and
  - (c) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a) or (b).

11. A vector comprising the DNA of Claim 4.
12. A host cell comprising the vector of Claim 11.
13. A process for producing a *S. epidermidis* polypeptide comprising expressing from the host cell of Claim 12 a polypeptide encoded by said DNA.
14. A process for producing a cell which expresses a *S. epidermidis* polypeptide comprising transforming or transfecting the cell with the vector of Claim 11 such that the cell expresses the polypeptide encoded by the DNA contained in the vector.
15. A polypeptide comprising an amino acid sequence which is at least 70% identical to a polypeptide set forth in the Sequence Listing.
16. A polypeptide comprising an amino acid sequence which is at least 80% identical to a polypeptide set forth in the Sequence Listing.
17. A polypeptide comprising an amino acid sequence which is at least 90% identical to a polypeptide set forth in the Sequence Listing.
18. A polypeptide comprising an amino acid sequence as set forth in the Sequence Listing.
19. An antibody against the polypeptide of claim 18.
20. An antagonist which reduces or inhibits the activity of the polypeptide of claim 18.
21. A method for the treatment of an individual having need to reduce or inhibit the activity of the polypeptide of Claim 18 comprising administering to the individual a therapeutically effective amount of the antagonist of Claim 20.
22. A complex of a polypeptide and a binding molecule which comprises the polypeptide of Claim 18 and a binding molecule that is capable of antagonising the activity of the polypeptide.
23. A process for diagnosing in a subject a disease related to expression of the polypeptide of claim 18 comprising detecting the presence in the subject of a nucleic acid sequence encoding said polypeptide.
24. A diagnostic process comprising detecting the presence of the polypeptide of claim 18 in a sample derived from a subject.
25. A method for identifying compounds capable of inhibiting the activity of the polypeptide of claim 18 comprising:
  - (a) contacting a cell expressing the polypeptide on the surface thereof with a selected compound, under conditions to permit binding to the polypeptide in the presence of a component capable of providing a detectable signal in response to the binding of the compound to said polypeptide; and

(b) detecting the presence or absence of a signal generated in response to the binding of the compound to the polypeptide,  
the presence of a signal indicating a compound capable of inhibiting the activity of the polypeptide.

26. A method for inducing an immunological response in a mammal which comprises inoculating the mammal with the polypeptide of Claim 15, or a fragment or variant thereof, adequate to protect said animal against infection from *S. epidermidis*.

27. A method of inducing an immunological response in a mammal which comprises delivering a gene encoding the polypeptide of Claim 15, or a fragment or variant thereof, and obtaining expression of the gene *in vivo* in order to induce an immunological response to produce antibody to protect said animal against infection from *S. epidermidis*.

28. An immunological composition comprising a DNA capable of expressing a polypeptide of Claim 15 which, when introduced into a mammal, induces an immunological response in the mammal, and a pharmaceutically acceptable carrier.

29. An immunological composition comprising a polypeptide of claim 15 which, when introduced into a mammal, induces an immunological response in the mammal, and a pharmaceutically acceptable carrier.